DOCUMENT RESUME

ED 278 678 TM 870 095

AUTHOR Hand, Carol A.; Prather, James E.
TITLE The Impact of Aggregation Bias upon the Interpretation of Test Scores across Schools.
PUB DATE Aug 86
NOTE 9p.; Paper presented at the Annual Meeting of the American Statistical Association (Chicago, IL, August, 1986).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Academic Achievement; *Analysis of Variance; *Cluster Grouping; College Entrance Examinations; Credits; Educational Background; Grade Point Average; Higher Education; Least Squares Statistics; Meta Analysis; Regression (Statistics); Scores; *Statistical Bias; *Test Interpretation; Undergraduate Students; Vectors (Mathematics)
IDENTIFIERS Scholastic Aptitude Test

ABSTRACT

This paper investigates the problems of the impact of grouping on measuring academic performance. It focuses on inferences made from the group level to the individual level. Data for a large state college system with over 30 individual colleges is used. The problem of aggregation bias is studied using the analysis of covariance and is related to the clustering approach of generalized least squares. To illustrate the question, actual data on academic performance at the individual and group level is explored. The analysis is done by gender and minority group status. It is found that the individual level relationships and the college level relationships were generally quite different, with regression coefficients often having different signs. Based on the results of this research it is concluded that it would be generally inappropriate to use grouped data to investigate academic performance across colleges. (JAZ)

THE IMPACT OF AGGREGATION BIAS UPON THE INTERPRETATION OF TEST SCORES ACROSS SCHOOLS

Carol A. Hand
James E. Prather
Office of Institutional Research
Georgia State University

Paper presented at the Annual Meeting of the
American Statistical Association
Chicago, Illinois
August, 1986

The Impact of Aggregation Bias Upon the Interpretation of Test Scores Across Schools

It is often necessary and convenient to use aggregate data concerning questions that are to a large extent about individuals. Research in public policy, education, demography, economics and sociology has often focused on the proper specification of models that rely on aggregate level data. The question of inferences that can be made about microlevel relationships from macrolevel data have been debated. The solution to this problem has been subject to continuous refinements of techniques designed to mitigate the problems of aggregate data.

This paper focuses on inferences made from the group level to the individual level. Since it is not possible to divorce substantive problems of model formation from methodological questions concerning technique, the example of making inferences about test scores across schools and colleges will be used to explore questions about aggregation bias.

Previous research on analyzing grouped or aggregate data has followed two separate paths of development. One perspective is represented by the seminal work of Robinson (1950) in sociology, while another is represented by the work of Prais and Aitchison (1954) in economics. Robinson's ecological correlation approach and the grouping data approach of Prais and Aitchison are complementary.

Analysis of Variance Approach to Aggregation

The analysis of covariance method to grouping grows out of Robinson's approach to aggregation. The analysis of covariance method is illustrated by partitioning the sums of squares about the mean for the dependent variable into explained sums of squares due to the covariates and to the groups along with the residual sums of squares.

Following the notation of Johnston (1972:192-207) a simple model is defined as

$$y = X + u \tag{1}$$

Where the sample y is a column vector (n x 1) of micro-level observations composed of p sub-vectors--i.e., the groups. The independent variables are the X matrix (n x k) divided into p groups and the first column is all ones to allow a constant term, while B is a vector (k x 1) of the estimators. The vector u contains stochastic noise values where E(u) = 0. To incorporate the possible effect of the p groups, then an expanded model is

$$y = D\alpha + XB + u, \tag{2}$$

which allows the p groups to have different constant terms, thus α is a vector of (p - 1) elements. The D' matrix is of dummy variables with order (Mp x [p-1]), where $M = \sum_{i=1}^{p} m_i$ is the sum of the number of observations in each p, for instance:

$$D' = \begin{vmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \end{vmatrix} \tag{3}$$

Remembering that D has p groups, with each p having m elements. To estimate (1) above, start with

$$y = X\hat{B} + s, \tag{4}$$

which can be estimated by

$$\hat{B} = (X'X)^{-1}X'y, \tag{5}$$

where s gives the least square residuals. An additional relationship may be derived as,

$$y'y = \hat{B}'X'y + s's \tag{6}$$

Returning to (2) above, the estimation of

$$y = D\hat{\alpha} + X\hat{B} + e \tag{7}$$

becomes

$$\begin{matrix} \hat{\alpha} \\ \hat{B} \end{matrix} = \begin{vmatrix} D'D & D'X \\ X'D & X'X \end{vmatrix}^{-1} \begin{matrix} D'y \\ X'y \end{matrix} \tag{8}$$

The Generalized Least-Squares Approach to Aggregation

In this section, if we start with (1) of the previous section the grouping of observations into p groups and taking means yields (... ston, pp. 228-241):

$$\bar{y} = \bar{X}\bar{B} = \bar{u} \tag{12}$$

Then the ungrouped data are related to the aggregated ... forms,

$$\bar{y} = Gy \tag{13}$$

$$\bar{X} = GX \tag{14}$$

$$\bar{u} = Gu \tag{15}$$

with G as the grouping matrix of (m x n). The form of G is for instance,

$$G = \begin{vmatrix} 1/1 & 1/1 & 0 & 0 & 0 & 0 & 0 & \ldots & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \ldots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & 1/p \end{vmatrix} \tag{16}$$

While $E(\bar{u}) = 0$ it is also noted that

$$E(\bar{u}\bar{u}') = \sigma^2 I \tag{17}$$

which means that the estimators will be unbiased but inefficient. However, it is the case that

$$E(\bar{u}\bar{u}') = \sigma^2 GG' \tag{18}$$

which is efficient. To estimated B, the generalized least squares is

$$b - [\bar{X}'(GG')^{-1}\bar{X}]^{-1}X'(GG')^{-1}\bar{y} \tag{19}$$

and

$$\mathrm{var}(b) = \sigma^2[\bar{X}'(GG')^{-1}\bar{X}]^{-1} \tag{20}$$

Here, generalized least squares overcomes the heteroscedastic problem (17) by inserting the grouping factor G in (18). The expression $(GG')^{-1}$ is actually a weighting matrix which contains the numbers in each group. Note that the generalized least squares estimates are not as efficient as the ungrouped ones.

The Effect of Aggregation on Interpreting Test Scores

Of substantive interest is the policy question of the relationship of test scores and academic background with performance in college level work.

The availability of grouped data by colleges, especially in terms of historical data, has made meta-analysis possible. This section seeks to explore the impact of grouping by college upon the relationship of college performance as measured by grade point average (GPA) with Scholastic Aptitude Test - Verbal (SAT-V) and Scholastic Aptitude Test - Mathematics (SAT-M) scores, and high school average (HSA). Data to investigate these questions were obtained for a large state college system with over 30 individual colleges. Over 45,000 undergraduate students who had one or more academic terms are included. As a control for college experience, the number of credit hours attempted and earned were employed.

Often when test scores are used as indicators of academic performance, questions of fairness to ethnic, racial, and gender groups have been raised. In this paper, the analysis of the influence of aggregation will be done by gender and minority groups status.

To test for the impact of aggregation on the relation of GPA with SAT scores, HSA, and credit hours, an analysis of variance was performed where the groups were the colleges and the test scores, HSA, and credit hours were the covariates. The partitioned sums of squares (SS) for this analysis are presented in Table 1. These show that the covariants are more strongly related to GPA for white females (WF) and white males (WM) than for black students of either sex. However, the R^2s for black females (BF) and black males (BM) are moderately strong for this type of data. The SS associated with the grouping, i.e. college effect shows that the college is more important for the BF and BM groups than for the WM and WF groups, accounting for more than 10% of the SS. Colleges account for 6% of the variables for the WF and WM groups. Clearly, colleges as a grouping factor have an independent influence on GPA even when controlling for test scores and academic background. College is an independent grouping factor in no sense like a random grouping.

Given that college as a grouping factor has a clear impact on GPA, it is interesting to review the regressions of the covariates with GPA to investigate if there is any aggregation bias present between the individual and college level. These regressions are presented in Table 2. As would be expected, the R^2 for the college level is somewhat higher than the individual level. Of the five regressions, the regression coefficient for SAT-V scores is negative at the college level in four cases. For WF, the coefficient for SAT-M is negative at the college level. These negative coefficients at the college level, when the individual level coefficients are positive, are strongly indicative of the bias introduced by grouping. The decline in magnitude of the HSA coefficients for four of the five regressions is an additional substantive outcome of the grouping factor. It is clear that the regression coefficients are markedly changed when group data is employed. Thus the typical measures of importance in the regression analysis were altered considerably by the aggregation effect. Anyone hoping to ascertain individual relationships from aggregate data of this type should expect little likelihood of success.

Summary and Discussion

The purpose of this paper has been to investigate the problem of the impact of grouping on measuring academic performance. This problem was explored using the analysis of covariance approach and is related to the clustering approach of generalized least squares. To illustrate this question, actual data on academic performance at the individual and grouped (college) level was explored. It was found that the individual level relationships and the college level relationships were generally quite different, with regression coefficients often having different signs. Based upon this research, it would seem to be generally inappropriate to use grouped data to investigate academic performance across colleges.

Table 1

Sums of Squares of GPA for Analysis of Covariance

Source of Variation	Sums of Squares									
	BF	R^2	BM	R^2	WF	R^2	WM	R^2	Total	R^2
Covariates	470	.17	287	.19	3379	.34	3178	.37	8093	.33
Colleges	327	.12	158	.10	576	.06	496	.06	1335	.05
Residual	1911		1085		6025		4977		14,988	
Total	2708		1530		9980		8651		24,416	
N	5931		3308		18,771		16,765		45,475	

Table 2

Regressions for Individuals and Colleges of GPA

Independent Variables	BF		BM		WF		WM		Total	
	Individual	Colleges	Individual	Colleges	Individual	Colleges	Individual	Colleges	Individual	C
SAT Verbal	.0008	-.0029	.0003	-.0043	.0019	.0056	.0012	-.0036	.0014	
SAT Math	.0006	.0016	.0006	.0028	.0004	-.0042	.0008	.0040	.0005	
H.S. Average	.2245	-.0233	.2312	.2647	.3709	.1355	.3066	.0017	.3497	
Hours Attempted	-.0059	.0007	-.0120	-.0005	-.0115	.0028	-.0168	-.0020	-.0121	
Hours Earned	.0081	.0053	.0143	.0029	.0131	-.0006	.0195	.0015	.0143	
Constant	.9840	2.1756	1.1164	1.6248	.3756	1.4998	.4937	1.9899	.5589	
Inverse College Size		.3215		.0563		1.8373		4.0176		1
N	5931	33	3308	32	18,771	33	16,765	33	45,475	
R^2	.17	.41	.19	.22	.34	.63	.37	.56	.33	
S.E.E.	$\pm$.61	$\pm$.18	$\pm$.61	$\pm$.22	$\pm$.59	$\pm$.16	$\pm$.57	$\pm$.16	$\pm$.60	